AMERICAN SOCIETY FOR MICROBIOLOGY | Antimicrobial Agents and Chemotherapy

Comprehensive Screening for Naturally Occurring Hepatitis C Virus Resistance to Direct-Acting Antivirals in the NS3, NS5A, and NS5B Genes in Worldwide Isolates of Viral Genotypes 1 to 6

Juan Ángel Patiño-Galindo,[a,d] Karina Salvatierra,[b,c*] Fernando González-Candelas,[a,d] F. Xavier López-Labrador[a,b,c,d]

Joint Units in Infection and Health[a] and Genomics and Health,[b] FISABIO-Public Health/Cavanilles Institute for Biodiversity and Evolutionary Biology, University of Valencia, Valencia, Spain; Virology Laboratory, Genomics and Health Area, FISABIO-Public Health, Generalitat Valenciana, Valencia, Spain[c]; CIBER-ESP (Centro de Investigación Biomédica en Epidemiología y Salud Publica), Instituto de Salud Carlos III, Madrid, Spain[d]

There is no comprehensive study available on the natural hepatitis C virus (HCV) polymorphism in sites associated with resistance including all viral genotypes which may present variable susceptibilities to particular direct-acting antivirals (DAAs). This study aimed to analyze the frequencies, genetic barriers, and evolutionary histories of naturally occurring resistance-associated variants (RAVs) in the six main HCV genotypes. A comprehensive analysis of up to 103 RAVs was performed in 2,901, 2,216, and 1,344 HCV isolates for the NS3, NS5A, and NS5B genes, respectively. We report significant intergenotypic differences in the frequencies of natural RAVs for these three HCV genes. In addition, we found a low genetic barrier for the generation of new RAVs, irrespective of the viral genotype. Furthermore, in 1,126 HCV genomes, including sequences spanning the three genes, haplotype analysis revealed a remarkably high frequency of viruses carrying more than one natural RAV to DAAs (53% of HCV-1a, 28.5% of HCV-1b, 67.1% of HCV-6, and 100% of genotype 2, 3, 4, and 5 haplotypes). With the exception of HCV-1a, the most prevalent haplotypes showed RAVs in at least two different viral genes. Finally, evolutionary analyses revealed that, while most natural RAVs appeared recently, others have been efficiently transmitted over time and cluster in well-supported clades. In summary, and despite the observed high efficacy of DAA-based regimens, we show that naturally occurring RAVs are common in all HCV genotypes and that there is an overall low genetic barrier for the selection of resistance mutations. There is a need for natural DAA resistance profiling specific for each HCV genotype.

Hepatitis C virus (HCV) infection is considered a major public health problem. More than 170 million people are chronically infected worldwide, with the consequent risk of developing liver diseases such as cirrhosis and liver cancer, which can eventually cause death (1, 2). HCV is a highly variable RNA virus which has been classified into 7 known genotypes (3). Genetic distances among genotypes reach up to 30% (4). Such diversity can be explained by an evolutionary rate of a magnitude of 10^{-3} substitutions per site and year (5). These differences at the genomic level also appear to be relevant at the clinical level. Treatment of chronic HCV infection with peginterferon-ribavirin combination therapy (P/R) shows variable sustained virological response (SVR) rates depending on the infecting HCV genotype (GT), with average SVR rates of 46%, 80%, 66%, and 60% for GTs 1, 2, 3, and 4, respectively (6, 7). Even within HCV GT1, a significant difference in SVR between subtypes 1a and 1b has been reported (8).

In recent years, the field of HCV therapy is blooming because of the clinical development of direct-acting antiviral drugs (DAAs) that are more effective than P/R (SVR up to >90%) and can be given in interferon (IFN)-free regimens with reduced toxicity (9, 10). DAAs that have advanced to clinical trials target three essential proteins for the HCV life cycle, encoded by the nonstructural (NS) protein genes: inhibitors of the NS3 serine protease (protease inhibitors [PIs]), inhibitors of the NS5A protein (NS5A inhibitors), and inhibitors of the NS5B RNA polymerase, either nucleos(t)idic (NI) or non-nucleos(t)idic (NNI) (11). More than 20 different DAA compounds targeting any of these proteins have been approved or are being investigated in advanced clinical trials. These regimens result in an increase in SVR rates to above 90% and reduce the duration of treatment to 12 weeks or less (12). Despite the high SVR rates obtained with these antivirals in HCV GT1 infection, the high variability and epidemic history of HCV may condition the real-world effectiveness of DAAs in a significant proportion of patients infected with other viral genotypes. Natural HCV variation is generated and transmitted over time even in the absence of DAAs. Indeed, naturally occurring DAA resistance-associated HCV variants (RAVs) have already been reported in DAA-naive patients, with some RAVs showing differential prevalence between genotypes and subtypes (13–21). A relevant example is the NS3-Q80K variant associated with resistance to simeprevir (SMV), which was previously found to be present in >30% of GT1a sequences but

Received 16 November 2015 **Returned for modification** 1 December 2015
Accepted 1 February 2016

Accepted manuscript posted online 8 February 2016

Citation Patiño-Galindo JÁ, Salvatierra K, González-Candelas F, López-Labrador FX. 2016. Comprehensive screening for naturally occurring hepatitis C virus resistance to direct-acting antivirals in the NS3, NS5A, and NS5B genes in worldwide isolates of viral genotypes 1 to 6. Antimicrob Agents Chemother 60:2402–2416. doi:10.1128/AAC.02776-15.

Address correspondence to F. Xavier López-Labrador, F.Xavier.Lopez@uv.es.

* Present address: Karina Salvatierra, Facultad de Ciencias Exactas, Químicas y Naturales, Universidad Nacional de Misiones, Posadas, Argentina.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/AAC.02776-15.

almost absent in GTs 1b, 2, 3, 4, and 5 (13, 16, 21). Current clinical guidelines include NS3-Q80K sequencing to evaluate treatment with SMV for GT1a-infected patients (9, 10). Other studies noted different rates of naturally occurring RAVs in NS5B. The NS5B-C316N variant seems highly prevalent in GT1b but not in GT1a isolates (17, 20), and NS5B-414L and -423I seem highly prevalent in GTs 4 and 5, respectively, but almost nonexistent in other HCV genotypes (20). Another relevant aspect rarely as yet explored is whether, for every resistance-associated site, differences in the genetic barrier (the minimal number of nucleotide mutations required to generate a given RAV) exist between HCV genotypes. This issue has been explored only for the emergence of substitutions causing drug resistance to some DAAs and for some particular, but not all, viral genotypes (20, 22).

To our knowledge, to date there is no study comprehensively describing the natural HCV variation and polymorphism in all three targets of approved DAAs (NS3, NS5A, and NS5B) and comparing the frequencies of natural RAVs between different viral genotypes and within isolates from the same genotype. Here, we report the analysis of an extensive data set of HCV sequences from DAA-naive patients available to date, trying to answer the following questions: (i) how prevalent natural RAVs are, (ii) whether there are inter- and/or intragenotypic differences in their frequencies, (iii) whether naturally occurring multiple resistance to different DAA classes exists, (iv) whether the genetic barrier for resistance differs between HCV genotypes, and (v) whether natural RAVs are generated *de novo* or are characteristic of a subset of viruses being already transmitted over time.

MATERIALS AND METHODS

HCV data sets. In April 2013, data mining was performed on three different databases: the European HCV database (EuHCVdb; http://euhcvdb.ibcp.fr/), the Los Alamos HCV database (LANL; http://hcv.lanl.gov/components/sequence/HCV/search/searchi.html), and The Virus Pathogen Database and Analysis Resource (VIPRBC; www.viprbrc.org/) in order to retrieve any nucleotide sequence including HCV nucleotide positions 3420 to 9000 (H77 reference, GenBank accession no. NC_004102) (23–25). A total of 2,931 sequences were retrieved: 1,328 sequences from LANL, 847 sequences from EuHCVdb, and 756 sequences from VIPRBC. A nonredundant data set of 1,407 sequences spanning the three NS3, NS5A, and NS5B genes was created using an in-house Perl script. To complement this data set with independent sequences from NS3, NS5A, and NS5B, in December 2013 additional searches were performed in LANL, EuHCVdb, and GenBank (www.ncbi.nlm.nih.gov/GenBank/) databases, expanding the initial data set to 4,551 NS3, 7,385 NS5A, and 2,182 NS5B nonredundant sequences. Low-quality sequences (those containing stop codons and/or with >1.2% of the amino acid positions consisting of indeterminate residues) and non-human-derived sequences (accession numbers were searched according to the keyword "HCV" and each of the following: "recombinant," "replicon," "mouse," "*Mus musculus*," "chimpanzee," "chimp," "*Pan troglodytes*," "replicon," "plasmid," "cell culture," "construct," and "chimera") were excluded. Only those HCV sequences derived from DAA-naive patients were retained, excluding all sequences from DAA-treated patients described in the literature up to January 2015. Reference isolates for the different HCV genotypes/subtypes were chosen in accordance with reference 3 and included in the data sets (available upon request).

Nucleotide sequence alignments were obtained with MUSCLE (26, 27) and manually edited with MEGA5.2 (28). In order to identify very closely related sequences (i.e., clones or sequential samples from the same individual), a clustering was performed for each data set using CD-HIT (29), setting the threshold of similarity to 0.95. Further exclusion of sequences originating from the same patient at different time points was done after retrieving additional information from GenBank.

The final data comprised four data sets: (i) NS3-NS5A-NS5B (concatenated set; 1,143 sequences), (ii) NS3 set (2,936 sequences), (iii) NS5A set (2,242 sequences), and (iv) NS5B set (1,376 sequences). All sequences from the concatenated data set were present in the corresponding individual gene data set. The genotype of the sequences was confirmed by using two different methods: (i) the CRP COMET subtyping tool (http://comet.retrovirology.lu/hcv/) and (ii) maximum likelihood (ML) phylogenetic analyses performed with PhyML using the GTR+GAMMA model (30). Recombinant and incorrectly genotyped sequences were excluded from the final data sets (n = 17, 35, 26, and 32 for NS3-NS5A-NS5B, NS3, NS5A, and NS5B sets, respectively).

Naturally occurring polymorphism at positions associated with resistance to DAA (natural RAVs). A compilation of 99 RAVs previously described in the literature in NS3, NS5A, and NS5B (31) (see Table S1 in the supplemental material) were used for computing the number and type of amino acid variants at the corresponding positions with BioEdit (32), first for HCV genotypes 1 to 6 and thereafter for subtypes 1a and 1b separately. To calculate the allele frequencies for each RAV, the BioEdit output table was further analyzed with R (33) at each position associated with resistance to DAA. Natural RAVs from all the analyzed positions associated with major resistance to approved DAAs were identified, according to their published phenotypic fold change in 50% inhibitory concentrations (IC_{50}s) of >10 on HCV replicon culture (see Table S1 in the supplemental material). Major mutations compiled in references 20 and 22 for which no >10-fold change in IC_{50} was obtained in our bibliographic search were also considered.

Amino acid signature patterns of resistance to DAA for HCV genotypes 1 to 6. Amino acid signature patterns for each position in the NS3, NS5A, and NS5B proteins associated with resistance to DAAs were inferred using VESPA (34), The HCV Con1 replicon or all GT1 amino acid sequences were used as background sets, which were compared with each of the query HCV genotypes/subtypes. For a given HCV GT, two analyses were performed: first to determine the consensus amino acid at each position (threshold = 0.50) and thereafter to detect the signature amino acids at each position (threshold = 1.00).

Searching for natural multiple resistance to different DAA classes. For each HCV genotype, the data set containing the concatenated NS3-NS5A-NS5B sequences was analyzed to determine the frequency of viral haplotypes with multiple RAVs. A data frame was created using in-house Perl and R scripts (33), with every row representing a sequence and each column representing a resistance-associated position. We computed the haplotype frequencies considering only resistance positions. The frequencies of resistant haplotypes for each analyzed HCV genotype/subtype were inferred with Arlequin v3.1 (35). Given the large number of possible haplotype combinations, the dimensions of the data frames were reduced by considering only 37 positions consistently associated with resistance to approved DAAs (see Table S2 in the supplemental material).

Calculation of the genetic barrier. The genetic barrier for the evolution of DAA-susceptible codons into RAVs was calculated according to a previous model applied to human immunodeficiency virus and HBV (36, 37) and recently to HCV (20). Given a position related to RAVs, each nonresistance codon present with a prevalence of >1% in the data sets was compared with all codons encoding a resistance amino acid, assigning a score of 1 to transitions and a score of 2.5 to transversions, because transitions occur 2.5 times more frequently than transversions (20, 36). The genetic barrier was calculated as the addition of the values assigned to the transitions and/or transversions needed to convert a nonresistance codon into a RAV. For each nonresistance codon, only the RAV codon with the minimal genetic barrier was chosen (the most parsimonious one), together with its minimal score. Due to the high number of possible RAVs in the HCV genes

analyzed, the genetic barrier was calculated only for the most clinically significant ones: NS3-36M/A, 54A/S, 55A, 80K/L/R, 155K, 156S/T/V, 168A/E/T/V/Y, and 170A; NS5A-28T, 30R/H, 31V/M, and 93H/C; and NS5B-159F, 282T, 316N/Y, 414T/V, 448H/C, 495L/Q/S, 554S, 556G, and 559G (reviewed in reference 38).

Evolutionary history of naturally occurring RAVs. For each data set, an ML phylogenetic tree was obtained with PhyML (30) with the GTR+GAMMA model for nucleotide substitutions and rooted with representative sequences from HCV genotypes 1 to 6. The evolutionary history of natural RAVs for each GT was analyzed with Mesquite (version 2.74; available at http://mesquiteproject.org/mesquite/download/download.html) using the phylogenetic tree obtained and the complete amino acid alignment of each GT as input. The evolution for each amino acid position of interest along the lineages of the phylogeny was inferred using a parsimony approach. Clades were defined as groups of clustered sequences with approximate likelihood ratio test (aLRT) values higher than 0.85.

RESULTS

Frequency of natural RAVs in HCV genotypes 1 to 6. A total of 2,901, 2,216, and 1,344 NS3, NS5A, and NS5B curated sequences, respectively, were included in the analyses of HCV GTs 1a, 1b, 2, 3, 4, 5, and 6. These data sets included 1,126 isolates with sequences spanning all of the complete NS3, NS5A, and NS5B genes.

Tables 1 to 4 (also see Table S3 in the supplemental material) summarize the amino acid frequencies at the most relevant resistance-associated positions in NS3, NS5A, and NS5B, by HCV genotype and the breakdown by subtypes. Among NS3 RAVs to approved protease inhibitors (PIs), V36L (low-level resistance to boceprevir [BOC]/telaprevir [TPV]/SMV/paritaprevir) was highly frequent in GT2 (99.3%), GT3 (100%), GT4 (100%), GT5 (100%), and GT6 (18.9%) but present in only 1.4% and 0.8% of GT1a and GT1b isolates, respectively. Variants in T54 associated with resistance to BOC/TPV showed low frequencies in GT1 (2.29% and 2.38% for GT1a and GT1b, respectively) and were rare in other GTs (found only in one GT2 and two GT3 isolates). V55A/I variants (associated with resistance to BOC) were concentrated in GT1a sequences (48 isolates, 3.91%). Only six non-GT1a isolates carried these variants (five GT1b and one GT3). The NS3-Q80K variant (resistance to SMV) was, as expected, highly frequent in GT1a (36.6%) but also in GT5 (100%) and GT6 (24.3%), whereas Q80G was a signature of GT2 (100% of isolates). R117H (resistance *in vitro* to BOC/TPV) was rare, present only in GT1 and with a very low prevalence (0.67% and 3.02% of GT1a and -1b sequences, respectively). Mutations at position NS3-122 (resistance to SMV) were common and differentially distributed among genotypes (122G/N was present in 7.24% of GT1a and in 13% of GT1b sequences; 122N in 33.8% of GT6 sequences; and 122T in 85.9% of GT4, 83.8% of GT5, and 48.6% of GT6 sequences, respectively). While NS3-122R was present only in GT2 (71.89%), 122A was unique to GT5 (10.8%). The major RAVs for PIs are located at NS3 positions 155, 156, and 168. R155K was very rare, present in only 13 GT1a, one GT1b, and two GT3 isolates (frequencies of 0.88%, 0.10%, and 1.71%, respectively). There was only one A156 variant in the data set: one GT6 isolate with A156V. In contrast, whereas D168 variants were uncommon in GT1 (0.20% and 0.91% of GT1a and GT1b sequences, respectively), D168Q was fixed in GT3 (100%), and 168E (resistance to most noncovalent PIs) was found in 48.6% of GT5 and 2.7% of GT6 sequences. Only three isolates (GTs 1a, 1b, and 6) showed V170A/T (resistance to BOC/SMV/asunaprevir), and I170 was dominant in GTs 1a, 2, 3, and 5 (97.3%, 84.61%, 98.52%, and 89.74%, respectively) and common in GTs 1b and 5 (30.94% and 31.08%, respectively). Finally, M175L (low-level resistance to BOC) was characteristic of HCV GTs 1a, 2 (99% of isolates), 3, 4, and 5 (100% of isolates) but present in only 1.4% of GT1b sequences.

For mutations occurring in NS5A, L28M (low-level resistance to daclatasvir [DCV]/ledipasvir [LDV]/ombitasvir) was present in 98.7% of GT3, 95.7% of GT1a, and 15% of GT4 sequences but in only 2% of GT1b sequences. On the other hand, L28V was common in GT6 isolates (43.8%) but rare in GT1a or GT4 sequences (3.3%). R30Q/L is a RAV studied in GT1b (Con1 replicon) which also causes resistance to DCV/LDV/ombitasvir. Although 30Q was found in only 4.8% of HCV-1b and 3.3% of GT3 sequences, it was present in 100% of GT5 sequences, and 30L was present in 28.3% of GT4 sequences. For variations in Q30 (Q30H/R/E) characterized in subtype 1a replicons, NS5A-30H was present in only 1.6% of GT1a sequences and 30R was present in 3.4% of GT2 sequences, but 30R was common for GT4 and GT6 (53.3% and 32.3%, respectively). Further, H58T (low-level resistance to LDV/ombitasvir) had a low prevalence of 1.2% in GT1b and 10% in GT4 sequences but of 40.6% in GT6 sequences. H58P (resistance in GT1a replicons) was present in 45.8% and 100% of GT6 and GT5 sequences, respectively, compared to only 2.7% of HCV-1a isolates. Finally, 153L (compensatory for resistance to thiazole analogues such as BP008 and DBPR110) was characteristic for GT5 (100%) and GT1a (96.9%) but less common in GT1b (18.1%), GT3 (12.4%), and GT4 (1.6%). The most important RAVs associated with clinical resistance to NS5A inhibitors are L31V/M and Y93H/N. L31M (low-level resistance to DCV/LDV) was highly represented in GT2 (81.2%) and GT4 (88.3%) but rare in GTs 1a, 1b, and 3 (0.9%, 4.6%, and 6.6%, respectively). Remarkably, within GT3, L31M variants were concentrated in non-GT3a isolates (see Table S3 in the supplemental material). NS5A-93H (resistance to DCV/LDV/ombitasvir) had a prevalence of 4.8% in GT1b, 3.1% in GT3, and 3.3% in GT4.

The prevalence of natural RAVs to NS5B NI and NNI was also differentially distributed among HCV genotypes. Among those variations affecting NIs, clinical resistance to sofosbuvir (SOF)/mericitabine (MCB) is associated with mutations in NS5B-S283, and L159F facilitates resistance (39, 40). S282T/R were present in just one isolate each for GTs 1a, 1b, 3, and 4 (frequencies of 0.17%, 0.24%, 1.24%, and 1.63%, respectively), and L159F was identified in only 4.5% of GT1b sequences. No other RAVs in these two NS5B positions were found for any genotype. In contrast to NIs, natural RAVs to NNIs were common and varied by virus subtype. Resistance patterns to NNIs depend on the binding site of the drug (38). NS5B-C316N/Y, M414T/V, Y448H/C, G554S, S556G, and D559G are associated with resistance to ABT-072 and dasabuvir, whereas A421V and P495L/S are associated with resistance to beclabuvir. Interestingly, C316N was found in up to 30% of GT1b and in 8.1% of GT4 sequences, all from non-GT4a isolates. It is worth noting that, while A421V represents a minority of GT1a (12.56%) or GT1b (5.03%) sequences, this variant is highly dominant in the other genotypes (80.32%, 98.77%, 91.94%, 37.5%, and 98.63% for GTs 2, 3, 4, 5, and 6, respectively). In contrast, only one sequence (GT6) carried P495 variants. At NS5B position 414, only 2.5% of GT1 sequences showed variants, while almost all GT2 (414Q/L, 97.5% and 2.5%, respectively) and GT4 (45.2%,

TABLE 1 Amino acid frequencies in HCV genotypes 1 to 6 at sites associated with resistance to NS3 inhibitors

NS3 position	Amino acid variant(s) (no. of isolates) for genotype[a]:						
	GT1a (n = 1,482)	GT1b (n = 992)	GT2 (n = 135)	GT3 (n = 117)	GT4 (64)	GT5 (n = 37)	GT6 (n = 74)
C16S	C (1,476), **S (1),** T (1)	C (984), T (2), **S (1),** Y (1)	A (111), T (18), **S (4),** P (2)	T (117)	T (64)	A (37)	T (74)
V36A/M/L/G/I/C	V (1,450), **L (21), M (7), I (1)**	V (978), **L (8), I (3)**	**L (134), M (1)**	**L (117)**	**L (64)**	**L (37)**	V (57), **L (14), I (3)**
A39V	A (1,461), T (8), G (7), S (3), **V (1)**	A (979), T (4), S (1)	**V (134),** I (1)	A (102), T (10), S (5)	A (60), T (4)	A (37)	A (72), T (1), D (1)
Q41R/K/P/H	Q (1,467), **H (12)**	Q (988), **H (2)**	Q (135)	Q (116)	Q (64)	Q (36), **H (1)**	Q (74)
F43S/C/Y/V/I/L	F (1,481), **Y (1)**	F (992)	F (135)	F (117)	F (63)	F (36)	F (74)
I48V	I (1,443), **V (33)**	I (382), **V (597),** L (4), F (2), T (1)	I (132), **V (3)**	I (15), **V (98),** L (4)	I (25), **V (39)**	I (37)	I (42), **V (28),** L (4)
T54A/S/V/G/C	T (1,446), **S (34)**	T (966), **S (23)**	T (133), **A (1)**	T (117)	T (59), **S (2)**	T (37)	T (73)
V55A/F/I/K/T	V (1,410), **I (36), A (32)**	V (984), **I (3), A (2)**	V (134), G (1)	V (116), **I (1)**	V (64)	V (36), L (1)	V (73)
D79E	D (1,478), **E (2),** N (1)	D (990), N (2)	D (1), **E (133),** K (1)	D (112), **E (4),** N (1)	D (64)	D (37)	D (71), **E (3)**
Q80K/L/N/R/H/G	Q (884), **K (542), L (28), R (9), N (7),** S (1), M (1)	Q (948), **L (33), R (5), K (4)**	**G (135)**	Q (116), **R (1)**	Q (63)	**K (37)**	Q (55), **K (18), L (1)**
A87T	A (1,435), S (22), **T (7),** C (5), I (1)	A (969), S (13), P (2), **T (2),** M (1), E (1)	A (1), S (127), **T (3),** N (3), G (1)	A (114), S (3)	A (56), S (5), C (1), V (1)	S (35), **T (2)**	A (54), S (19), **T (1)**
Y105C	Y (1,476), F (3)	Y (976), F (11)	Y (135)	Y (117)	Y (30), F (34)	Y (37)	Y (70), F (3), **C (1)**
R109K	R (1,482)	R (986), M (1), **K (1)**	R (135)	R (117)	R (63)	R (37)	R (73)
R117H	R (1,448), C (19), **H (10),** S (2), L (1)	R (929), **H (30),** C (28), Q (1), F (1)	R (135)	R (117)	R (64)	R (37)	R (74)
S122G/A/R/N/T	S (1,365), **G (96), N (11), T (2),** C (2)	S (798), **G (94), T (56), N (35),** C (2)	**R (110),** K (23), **T (1)**	S (117)	S (5), **T (55), N (4)**	S (2), **T (31), A (4)**	S (13), **T (36), N (25)**
R123H/K	R (1,477), **K (3)**	R (983), **K (8)**	R (135)	T (117)	R (61), **K (1)**	R (34), **H (1)**	R (74)
I132V	I (1,466), **V (5),** L (2)	I (282), **V (692),** L (10)	I (7), L (127)	I (8), L (109)	I (59), L (4), **V (1)**	I (35), L (1), **V (1)**	I (51), L (22)
S138T/D/P	S (1,479)	S (988), F (1)	S (135)	S (117)	S (63), F (1)	S (37)	S (71), F (1)
R155K/T/I/M/G/L/S/Q/P/N/W	R (1,463), **K (13)**	R (991), **P (1)**	R (135)	R (115), **K (2)**	R (63)	R (36)	R (73)
A156S/T/V/I/F/N/G/D	A (1,478)	A (992)	A (134)	A (117)	A (64)	A (37)	A (73), **V (1)**
V158I/M	V (1,480)	V (990), **I (1)**	V (130), **I (3), M (2)**	V (115), A (1), **I (1)**	V (62), L (1), **I (1)**	V (33), **M (2),** L (1)	V (71), **I (2)**
V163L	V (1,478)	V (984), I (4)	V (133), I (1), A (1)	V (114), E (1)	V (62)	V (36), A (1)	V (74)
D168Q/A/Y/V/E/T/N/P/I/H/G/F/S/K	D (1,471), **E (3), G (1)**	D (979), **E (9)**	D (135)	**Q (116)**	D (64)	D (19), **E (18)**	D (72), **E (2)**
V170A/T/G/L/M	V (55), I (1,422), **T (1)**	V (682), I (307), **T (1)**	V (2), I (133)	V (11), I (105)	V (63), I (1)	V (3), I (33)	V (50), I (23), **A (1)**
E173G	E (1,481)	E (992)	E (133), D (1), **G (1)**	E (116), D (1)	E (63)	E (36)	E (74)
S174F/P	S (584), N (827), G (59), D (2), A (1), C (1)	S (960), A (19), **F (6),** T (4), H (1), L (1), C (1)	S (106), T (18), A (11)	S (3), T (106), A (7), I (1)	S (59), A (3)	S (4), N (32)	S (27), N (44), A (2), G (1)
M175L	**L (1,476)**	M (977), **L (14)**	**L (134),** I (1)	**L (117)**	**L (64)**	**L (37)**	M (74)
E176G	E (1,466), **G (7),** D (4), A (1), S (1)	E (986), Q (2), D (1), **G (1)**	D (110), A (17), N (8)	S (93), N (18), A (5), H (1)	E (62), A (1)	E (37)	E (46), Q (26), D (2)

[a] RAVs are indicated in bold.

TABLE 2 Amino acid frequencies in HCV genotypes 1 to 6 at sites associated with resistance to NS5a inhibitors

NS5a position	Amino acid variant(s) (no. of isolates) for genotype[a]: GT1a (n = 861)	GT1b (n = 810)	GT2 (n = 149)	GT3 (n = 226)	GT4 (n = 60)	GT5 (n = 14)	GT6 (n = 96)
Reference GT1b (Con1)							
Q24L	Q (2), K (844), R (5), E (4), S (1)	Q (788), K (11), R (10)	S (113), T (35), A (1)	S (218), **L (5),** A (1), G (1), T (1)	K (60)	Q (14)	K (65), Q (29), R (2)
L28M/V	L (1), **M (824), V (28),** T (6), I (2)	L (789), **M (16), V (3),** P (1)	L (90), F (56), C (3)	L (1), **M (223),** I (2)	L (48), **M (9), V (2),** I (1)	L (14)	L (22), **V (42),** F (11), T (11), A (5), **M (3),** G (1)
R30Q/L	R (4), **Q (840),** H (14), **L (2)**	R (756), **Q (39),** K (7), **L (3),** M (2), H (1)	R (5), K (143)	R (2), A (180), K (30), T (6), **L (3),** S (2), V (2), M (1)	R (32), **L (17),** S (4), T (4), **Q (2),** A (1)	**Q (14)**	R (3 1), S (41), A (22), N (1)
L31M/V/F	L (853), **M (8)**	L (766), **M (37),** I (3), **V (1), F (1),** P (1)	L (28), **M (121)**	L (208), **M (15), V (2)**	L (7), **M (53)**	L (14)	L (95), I (1)
P32L	P (860)	P 810	P (149)	P (226)	P (70)	P (14)	P (96)
Q54H/N/L/Y	**H (847), Y (11),** C (1), F (1)	Q (524), **H (253), N (9), L (6), Y (15),** C (1), E (1), T (1)	T (147), **N (1),** S (1)	S (196), T (29)	**H (60)**	S (13), **Y (1)**	**H (69),** R (10), **Y (7),** T (3), C (2), **N (2),** S (1)
P58S/A/L/T/H	P (23), **H (823),** R (6), Y (6), Q (1), C (1), D (1)	P (750), **S (30),** T (10), **L (7),** Q (6), **A (6)**	P (141), **S (6), H (1), T (1)**	P (215), **S (8),** R (1), **A (1), T (1)**	P (52), **T (6),** R (1)	P (14)	P (44), **T (39),** G (5), **S (4), A (2), L (1)**
Q62E/R/A/P/S	**E (830),** D (24), G (3), **A (1)**	Q (774), **E (16), R (6),** H (5), N (2), K (2), **S (1),** D (1), G (1), L (1)	N (128), **A (6),** S (4), T (4), V (3), H (1), I (1), **E (1)**	Q (1), **S (135),** T (46), M (7), D (7), L (6), **E (6),** V (5), **A (4), P (4),** I (4), N (1)	Q (4), **E (42), S (4),** N (4), D (3), **R (2),** V (1)	T (13), **A (1)**	Q (4), V (34), D (15), N (11), T (8), **A (8), S (5),** M (4), **E (3),** H (1), I (1), **R (1),** K (1)
A92T	A (854), P (7)	A (783), **T (16),** V (7), G (1)	A (1), C (142), S (6)	E (224), G (1)	A (57), **T (1)**	A (14)	A (96), P (1)
Y93H/N/C	Y (849), **H (5), C (4), N (1),** F (1)	Y (769), **H (39), C (2)**	Y (148), F (1)	Y (219), **H (7)**	Y (55), **H (2),** T (1), R (1), S (1)	T (14)	T (79), S (14), I (2)
F149L	F (858), **L (1)**	F (809)	F (149)	F (226)	F (58), **L (1),** S (1)	F (14)	F (96)
V153M/L/I	V (15), **L (834), I (4),** P (1)	V (657), **L (147), I (5),** T (1)	V (149)	V (197), **L (28)**	V (57), **L (1)**	**L (14)**	V (92), **I (4)**
M202L	M (856), **L (2),** Q (2), I (1)	M (807), **L (3)**	M (148), V (1)	M (225), **L (1)**	M (60)	M (14)	M (94), **L (2)**
M265V/T	M (858), **V (2),** L (1)	M (791), K (9), **V (5), T (1),** W (1)	M (148), L (1)	M (222), **V (2),** I (1), L (1)	M (13)	M (14)	M (57), **V (32),** L (3), A (2), I (1)
D320E	D (857), **E (2),** N (1), G (1)	D (796), **E (6),** S (3), N (2)	D (131), **E (7),** N (4), S (3), G (2), Q (1)	D (226)	D (43), **E (11),** A (4), S (1)	D (14)	D (94), N (1), G (1)
Y321N	Y (856)	Y (805), H (1)	Y (147)	Y (226)	Y (58)	Y (14)	Y (95), **N (1)**
Reference GT1a (H77)							
L23F	L (856), M (1)	L (807), P (2), I (1)	L (149)	L (220), P (5), I (1)	L (57)	L (14)	L (96)
M28T	M (824), V (28), **T (6),** I (2), L (1)	M (16), L (789), V (3), P (1)	L (90), F (56), C (3)	M (223), I (2), L (1)	M (9), L (48), V (2), I (1)	L (14)	M (3), V (42), L (22), F (11), **T (11),** A (5), G (1)
Q30H/R/E	Q (840), **H (14), R (4),** L (2)	Q (39), **R (755),** K (7), L (3), M (2), **H (1)**	K (143), **R (5)**	A (180), K (30), T (6), L (3), V (2), S (2), **R (2),** M (1)	Q (2), **R (32),** L (17), S (4), T (4), A (1)	Q (14)	S (41), **R (31),** A (22), N (1)
L31V/M	L (853), **M (8)**	L (766), **M (37),** I (3), **V (1),** F (1), P (1)	L (28), **M (121)**	L (208), **M (15), V (2)**	L (7), **M (53)**	L (14)	L (95), I (1)
P32L	P (860)	P (810)	P (149)	P (226)	P (60)	P (14)	P (96)
H58D/P	H (823), **P (23),** R (6), Y (6), Q (1), **D (1),** C (1)	**P (750),** S (30), T (10), L (7), Q (6), A (6)	H (1), **P (141),** S (6), T (1)	**P (215),** S (8), R (1), A (1), T (1)	**P (52),** T (6), R (1)	**P (14)**	**P (44),** T (39), G (5), S (4), A (2), L (1)
Y93N/C	Y (849), H (5), **C (4), N (1),** F (1)	Y (769), H (39), **C (2)**	Y (148), F (1)	Y (219), H (7)	Y (55), H (2), T (1), R (1), S (1)	T (14)	T (79), S (14), I (2)
D320E	D (857), **E (2),** N (1), G (1)	D (796), **E (6),** S (3), N (2)	D (131), **E (7),** N (4), S (3), G (2), Q (1)	D (226)	D (43), **E (11),** A (4), S (1)	D (14)	D (94), N (1), G (1)

[a] RAVs are indicated in bold.

TABLE 3 Amino acid frequencies in HCV genotypes 1 to 6 at sites associated with resistance to NS5B NIs

NS5B NI resistance position	Amino acid variant(s) (no. of isolates) for genotype[a]:						
	GT1a (n = 581)	GT1b (n = 417)	GT2 (n = 122)	GT3 (n = 81)	GT4 (n = 62)	GT5 (n = 8)	GT6 (n = 73)
A15G	A (562), S (14), T (3), V (2)	A (379), S (32), T (3), V (2), **G (1)**	A (4), **G (80),** S (32), C (2), R (1), I (1)	A (2), S (75), T (2), N (2)	A (62)	S (8)	A (71), V (2)
K72M	K (579), **M (1)**	K (416), N (1)	K (122)	K (80), R (1)	K (61)	K (6), R (2)	K (72)
S96T	S (581)	S (417)	S (122)	S (81)	S (62)	S (8)	S (73)
N142T	N (581)	N (385), S (29), **T (1)**	N (120), **T (2)**	N (78), S (3)	N (58), S (3), **T (1)**	N (8)	N (73)
L159F	L (581)	L (397), **F (19)**	L (122)	L (81)	L (62)	L (8)	L (72)
R222Q	R (579), D (1)	R (417)	R (122)	R (81)	R (62)	R (8)	R (65), K (5), S (3)
C223H/Y	C (581)	C (416), R (10)	C (121), W (1)	C (81)	C (62)	C (8)	C (73)
I239V/L	I (580), **V (1)**	I (416)	I (120), **L (2)**	I (78), **V (3)**	I (10), **V (52)**	I (8)	I (62), **V (10)**
S282T/R	S (580), **R (1)**	S (416), **T (1)**	S (122)	S (79), **R (1)**	S (61), **T (1)**	S (8)	S (73)
A300T	Q (296), R (260), K (21), L (4)	A (18), S (300), **T (95),** V (1), R (1), F (1)	L (105), R (8), Q (8), K (1)	**T (67),** S (6), K (5), M (2), R (1)	**T (46),** S (8), V (5), N (1), M (1)	L (8)	A (1), Q (53), E (5), K (4), M (3), S (3), **T (2),** N 1
L320I/F	L (579)	L (415), R (1)	L (122)	L (80), Q (1)	L (61)	L (8)	L (73)
V321I	V (580), **I (1)**	V (414), **I (3)**	V (120), F (1), **I (1)**	V (81)	V (59), **I (3)**	V (8)	V (73)
A396G	A (579)	A (417)	A (122)	A (81)	A (61), V (1)	A (8)	A (73)
Y586C	Y (544), **C (2)**	Y (399), N (1), F (1), **C (1)**	F (114), L (2), V (1)	F (57), V (1)	F (49)	F (7)	F (71), I (1)

[a] RAVs are indicated in bold.

37.1%, 16.12%, and 1.6% with variants 414L/V/I/Q, respectively) isolates did. Although NS5B-556 RAVs were rare for GT1 (556G/N in 7.2% and 1.4% of HCV-1b and 1a isolates, respectively), 556G was dominant in all non-1 HCV genotypes except GT6 (95.9%, 93.5%, 82.2%, 100%, and 4.2% of GT2, GT3, GT4, GT5, and GT6 sequences, respectively). In contrast, G554S was uncommon (7.4% and 2.5% of GT2 and GT3 isolates, respectively). Other natural RAVs to NNIs identified included L392I (resistance to deleobuvir and TMC647055), which was more frequently found in GT2 (69.7%) than in HCV-1b and GT6 (3.8% and 1.6%, respectively). Finally, RAVs P495L/Q/S (resistance to deleobuvir and TMC647055) were identified in only one GT6 isolate (P495L), and Y448H/C (resistance to tegobuvir) was identified in one GT1b isolate (Y448C).

Amino acid signature patterns and RAVs. We used the VESPA program to determine the overall resistance profile to DAAs within and between HCV genotypes and to identify the predominant and signature amino acids in positions associated with resistance at the three NS3, NS5A, and NS5B genes. Tables 5 to 8 (see also Table S4 in the supplemental material) show the consensus amino acid pattern for each HCV genotype and the breakdown by subtypes. For a given position, one amino acid was considered a signature when present in 100% of the sequences from a given genotype. For HCV GTs 2, 3, 4, and 5, all the sequences presented more than one natural RAV as a signature amino acid. For GT2, signature RAVs were NS3-Q80G and NS5B-I424V, C445F, I482L, and V494A. For GT3, signature RAVs were NS3-V36L, D168Q, and M175L and NS5B-I424V, C445F, and I482L. For GT4, two NS3, one NS5A, and three NS5B RAVs were found as signature amino acids (NS3-V36L and M175L, NS5a Q54H and NS5B-L419I, I482L, and R531K). In GT5, signature RAVs were NS3-V36L, Q80K, and M175L and NS5B-I363V, M423I, I424V, C445F, R531K, and S556G. For GT6, three signature RAVs were found in NS5B (NS5B-C455F, I482L, and V499A). Finally, we found several positions with differential signature/majority amino acid patterns between subtypes of the same GTs (Tables 5 to 8; see also Table S4). These positions were NS3-48, 132, 170, 174, and 175, NS5A-24, 28, 30, 54, 58, 62, and 153, and NS5B-19, 71, 300, 338, and 499 for GT1; NS3-122, 173, and 174, NS5A-24 and 28, and NS5B-15, 71, 338, and 392 for GT2; NS3-39, 132, and 174, NS5A-30, 31, 54, 62, and 153, and NS5B-300, 442, and 494 for GT3; NS3-105 for GT4; and NS3-48, 80, 87, 174, and 176, NS5A-24, 28, 30, 58, 62, and 265, and NS5B-300, 494, 486, 555, 556, and 571 for GT6.

Haplotypes with multiple RAVs to different DAA classes. For the subset of HCV isolates for which all the NS3, NS5A, and NS5B gene sequences were available (n = 1,126), the frequency of isolates carrying more than one RAV was calculated considering all positions associated with resistance. For each HCV genotype, the haplotypes with multiple RAVs are detailed in Table S5 in the supplemental material, including their frequencies. Haplotypes with more than one RAV accounted for 53% of GT1a sequences, with NS3-(Q80K+175L) being present in 32% of them. For GT1b, the frequency of sequences with more than one RAV was 28.5%, and 4.5% of GT1b sequences presented the haplotype NS3-(122G)+NS5B-(316N). It is worth noting that all (100%) the haplotypes of genotypes 2, 3, 4, and 5 carried more than one natural RAV. In GT2, the most prevalent haplotype with multiple RAVs (40%) was NS3-(36L+80G+122R+175L)+NS5A-(31M)+NS5B-(392I+414Q+556G). For GT3, NS3-(36L+168Q+175L)+NS5A-(28M) and NS3-(36L+168Q+175L)+NS5A-(28M)+NS5B-(556G) were present in 42.9% and 28.6% of the sequences, respectively. For GT4, there were two equally frequent haplotypes, NS3-(36L+122T+175L)+NS5A-(28M+31M)+NS5B-(414L+556G) and NS3-(36L+122T+175L)+NS5A-(31M)+NS5B-(414L+556G), both present in only 8.8% of the sequences. Up to 50% of the GT5 sequences showed the haplotype NS3-(36L+80K+122T+168E+170V+175L)+NS5B-(556G). Finally, for GT6, 67.1% of the

TABLE 4 Amino acid frequencies in HCV genotypes 1 to 6 at sites associated with resistance to NS5B NNIs

NS5B NNI resistance position	Amino acid variant(s) (no. of isolates) for genotype[a]:						
	GT1a (n = 581)	GT1b (n = 417)	GT2 (n = 122)	GT3 (n = 81)	GT4 (n = 62)	GT5 (n = 8)	GT6 (n = 73)
T19S/P	Q (574), E (5), D (1), **P (1)**	T (43), **S (368),** N (7), G (1)	E (118), G (2), D (1), R (1)	E (81)	T (17), **S (37),** A (4), **P (3),** N (1)	E (8)	E (73)
K50R	K (576), **R (5)**	K (403), **R (12),** Q (1)	K (122)	K (78), **R (3)**	K (61)	K (8)	K (69), **R (4)**
D55E	D (578), **E (3)**	D (517)	D (122)	D (81)	D (62)	D (8)	D (68), **E (2),** H (1), G (1)
M71V	**V (578),** I (2), A (1)	M (393), I (13), **V (10),** L (1)	**V (75),** I (47)	**V (71),** I (9)	I (59), T (1), **V (1)**	M (8)	M (3), **V (18),** I (45), T (6)
H95Q/R	H (575)	H (414), L (1), **Q (1)**	H (122)	H (78), N (2), K (1)	H (62)	H (8)	H (71), C (1), **R (1)**
V138I	V (1), **I (579),** L (1)	V (27), **I (390)**	V (10), **I (111),** A (1)	V (1), **I (80)**	**I (62)**	**I (8)**	V (2), **I (70)**
L314F	L (577), P (1), **F (1)**	L (417)	L (122)	L (81)	L (62)	L (8)	L (73)
C316Y/F/N/S	C (580)	C (287), **N (126),** H (1), R (1), **Y (1), S (1)**	C (121), W (1)	C (79), G (2)	C (54), **N (5),** H (3)	C (8)	C (73)
A338V	A (579)	A (31), **V (386)**	A (96), **V (24),** T (1)	A (75), **V (5),** T (1)	A (58), **V (3),** T (1)	A (8)	A (70), **V (1)**
I363V	I (578), **V (1)**	I (416), **V (1)**	I (122)	I (81)	I (62)	**V (8)**	I (73)
S365T/A/L/O/F	S (579)	S (416), **A (1)**	S (122)	S (81)	S (62)	S (8)	S (73)
S368A/T	S (577)	S (416), P (1)	S (122)	S (81)	S (62)	S (8)	S (70), **A (1)**
T389S/A	T (572), **A (3), S (2),** M (1), V (1)	T (400), **A (9),** I (4), **S (2),** M (2)	T (119), **A (1),** I (1), **S (1)**	T (58), E (22), D (1)	E (62)	Q (8)	T (52), V (13), I (5), L (1), **S (1), A (1)**
L392I	L (568), F (11), **I (3)**	L (396), **I (16),** F (4)	L (35), **I (85),** F (1)	L (81)	L (62)	L (8)	L (71), **I (1)**
N411S	N (578)	N (417)	N (120), T (2)	N (79), **S (2)**	N (62)	N (8)	N (73)
M414L/T/I/V/Q	M (577), **L (1), V (1)**	M (415), **L (1), I (1)**	**Q (119), L (3)**	M (81)	**L (28), V (23), I (10), Q (1)**	M (8)	M (71), A (1)
L419M/V/S/I	L (578)	L (411), **I (5)**	L (2), **I (117), V (3)**	L (3), **I (78)**	**I (61)**	L (8)	L (1), **I (72)**
A421V	A (505), **V (73),** M (1)	A (396), **V (21)**	A (23), **V (98)**	A (1), **V (80)**	A (5), **V (57)**	A (5), **V (3)**	A (1), **V (72)**
R422K	R (576), **K (3)**	R (417)	R (122)	R (81)	R (62)	R (8)	R (73)
M423T/V/I/A	M (561), **I (13), V (3), T (1), A (1)**	M (423)	M (122)	M (79), **I (2)**	M (62)	**I (8)**	M (73)
I424V	I (560), **V (19)**	I (381), **V (35)**	**V (122)**	**V (81)**	I (16), **V (54)**	**V (8)**	I (1), **V (72)**
M426T/V/I	M (531), L (41), F (1), **T (1), I (1)**	M (405), L (10), A (1), **T (1)**	M (120), L (2)	M (75), L (4), **V (1), I (1)**	M (61)	**T (1)**	M (63), C (10)
A442T	A (579)	A (405), **T (10),** V (2)	N (101), D (19), S (2)	A (23), P (54), S (4)	A (62)	A (7), **T (1)**	A (49), V (19), P (3)
C445F	C (573), G (1), Y (1)	C (411), **F (4),** W (1)	**F (122)**	**F (81)**	**F (62)**	**F (8)**	**F (73)**
I447F	I (578), **F (1)**	I (447)	M (118), L (4)	M (81)	M (62)	M (8)	I (39), M (32), L (1)
Y448H/C	Y (579)	Y (416), **C (1)**	Y (122)	Y (81)	Y (62)	Y (8)	Y (73)
C451R	C (572), Y (4), H (2), W (1)	C (333), T (46), Y (13), H (7), I (7), N (3), S (3), V (2), **R (1)**	V (120), A (1), I (1)	T (80), V (1)	T (62)	V (8)	T (73)
Y452H	Y (577), **H (2)**	Y (410), **H (7)**	Y (122)	Y (81)	Y (62)	Y (8)	Y (72), **H (1)**
C455F	E (567), Q (6), K (5)	E (412), Q (4), G (1)	C (1), N (87), S (30), T (4)	T (81)	T (62)	T (8)	C (1), T (43), S (24), N (3), A (2)

M/I462T	I (578), L (1)	I (416), V (1)	I (119), V (3)	I (80)	I (62)	R (8)	I (72), V (1)
R465G	R (575), K (1)	R (416), L (1)	R (122)	R (80)	R (58), K (4)	R (8)	R (73)
I482L/V/T/S	I (564), **L (1)**	I (411), **L (1)**	**L (119)**	**L (75)**	**L (60)**	I (8)	**L (72)**
A486V/I/T/M	A (563), D (1)	A (411)	A (119)	A (73), S (2)	A (55), S (1)	A (7), S (19)	A (34), G (38)
V494A/I	V (562), **I (1)**	V (411), **A (1)**	**A (119)**	C (56), **I (14), A (3),** S (2)	V (56), **A (2)**	V (8)	V (14), **A (57),** M (1)
P495L/A/S/T/Q	P (564)	P (412)	P (119)	P (75)	P (57)	P (8)	P (71), **L (1)**
P496A/T/S	P (563)	P (412)	P (118), **S (1)**	P (75)	P (58)	P (8)	P (72)
V499A	V (4), **A (545),** T (15)	V (364), **A (41),** T (6), I (1)	V (9), **A (106),** T (3), M (1)	V (1), **A (74)**	V (1), **A (57)**	**A (8)**	**A (72)**
R531K	R (467), K (84)	R (275), **K (136)**	R (15), **K (102)**	R (59)	**K (53)**	**K (8)**	R (32), **K (40)**
G554D/S	G (548)	G (402), **D (1)**	G (109), **S (9)**	G (57), **S (2)**	G (52)	G (8)	G (72)
Y555C	Y (548)	Y (411)	A (111), S (7)	V (51), I (4), A (3), G (1)	A (44), S (4), G (3)	A (8)	Y (34), G (19), F (17), V (1)
S556G/N/C	S (543), **G (4), N (1)**	S (373), **G (30), N (6),** D (1)	**G (117), C (1)**	S (1), **G (58)**	**G (49), N (2),** A (1)	**G (8)**	S (41), D (18), R (11), **G (2)**
G558R	G (548)	G (405)	G (114), S (4)	G (3), N (55), S (1)	G (51), E (1)	G (8)	G (70), A (1), V (1)
D559G/S/N	D (548)	D (403), **N (1)**	D (117), H (1)	D (59)	D (52)	D (8)	D (71), **G (1)**
W571R	W (545), **R (2)**	W (402), L (1), **R (1)**	L (101), S (6), F (4), Y (3), I (3)	H (39), Y (16), L (1), Q (1)	Y (49), F (1), H (1)	Y (5), N (1), C (1)	M (24), L (21), I (12), T (6), V (3), F (3), N (1)

[a] RAVs are indicated in bold.

sequences presented haplotypes with at least two natural RAVs, NS3-(Q80K+122N)+NS5A-(58T) and NS3-(36L+122T) (17.8% and 8.8% of the total haplotypes, respectively).

Genetic barriers. The genetic barrier for the evolution of DAA resistance was calculated for the analyzed data sets of each gene and HCV genotype. For each original codon with no resistance, the minimal genetic barrier was defined as the lowest score obtained from the total number of transitions (score = 1) and/or transversions (score = 2.5) needed to generate a given RAV. Genetic barrier scores were classified as "intermediate" for values lower than 5.0 but equal to or higher than 3.0, while values equal to or higher than 5.0 were classified as "high." Thus, both intermediate and high genetic barrier scores necessarily imply at least two nucleotide changes. The results are shown in Table S6 in the supplemental material. Considering the most frequent nonresistant codon for each position and GT, four positions in NS3, two in NS5A, and seven in NS5B presented very low genetic barriers. For NS3, the following codons had minimal scores of 1.0: V36(GTG) in GT1a; V55(GTC) in GTs 1a, 1b, 5, and 6; V55(GTT) in GT3; V55(GTA) in GT2; R155(AGA) in GT2; R155(AGG) in GTs 1a, 3, and 5; and V170(GTG) in GTs 4 and 6 or V170(GTA) in GT1b. Only for positions NS3-L36(CTT) in GT3, R155(CGC) in GTs 4 and 6, and Q168(CAG) in GT3 was a high genetic barrier evident (L36M score = 5; R155K score = 6; and Q168A/E/V score = 5), and not for all possible RAVs. Thus, most positions presented low or intermediate genetic barrier scores. A similar pattern was found among less frequent codons in all genotypes. For NS5A, two positions presented very low genetic barrier scores: M28(ATG) in GTs 1a and 3 and Y93(TAC) in GTs 1a, 1b, 2, 3, and 4. On the other hand, NS5A-T93(ACA) in GT5 and T93(ACC) in GT6 presented high genetic barriers toward all possible RAVs (all minimal scores were ≥5). Most codon variants found at lower frequencies also showed low or intermediate genetic barriers for all GTs. For NS5B, the following codons presented low genetic barrier scores for all possible RAVs: L159(CTT) in GT1b and L159(CTC) in GTs 1a, 2, and 3; L320(CTT) in GTs 1b and 5; M414(ATG) in GTs 1a, 1b, 3, 5, and 6; Y448(TAC) in GTs 1 to 6; Y448(TAT) in GT5; G554(GGT) in GT1b; G554(GGC) in GTs 1a, 2, 3, 4, 5, and 6; S556(AGC) in GTs 1a, 1b, and 6; and A559(GAC) in HCV GTs 1 to 6. In all HCV GTs, the generation of S282T was associated with a low genetic barrier score, requiring only one substitution but at the second base of the codon (see Table S4 in the supplemental material). Other highly frequent codons, such as C316(TGT) in GTs 1a, 2, and 5; P495(CCT) in GTs 2 and 6; and P495(CCC) in GTs 3, 4, and 5, presented high genetic barriers for change to some RAVs but low barriers for others. A high genetic barrier (score = 5) was found only for Q414(CAA), a polymorphism particularly frequent in genotype 2 isolates. Overall, most of the codon variants present at lower frequencies also showed low or intermediate genetic barriers to generate RAVs in all HCV GTs.

Finally, when analyzing each HCV genotype/subtype independently, the most frequent nonresistant codons from NS3, NS5A, and NS5B presented, on average, low genetic barrier scores. However, several differences in genetic barriers were observed between genotypes/subtypes. For NS3-36, the genetic barrier is very low in GT1a (V36A/M score = 1) but high in GT3 (L30A, score = 3.5; L36M, score = 5) isolates. Similarly, for NS3-R155, the genetic barrier is very low for GTs 1a (as previously known), 2, 3, and 5 (score = 1) but high for GTs 1b, 4, and 6 (score = 6). For NS5A

TABLE 5 Majority amino acids and signatures of HCV genotypes 1 to 6 at sites associated with resistance to NS3 inhibitors

NS3 position	Amino acid for sequence or genotype[a]: Con1 reference (1b)	GT1 consensus amino acid	All GT1	1a	1b	2	3	4	5	6
C16S	C	C	●	●	●	A	T*	T*	A*	T*
V36A/M/L/G/I/C	V	V	●	●	●	L	L*	L*	L*	●
A39V	A	A	●	●	●	**V**	●	●	●	●
Q41R/K/P/H	Q	Q	●	●	●	●	●	●	●	●
F43S/C/Y/V/I/L	F	F	●	●	●	●	●	●	●	●
I48V	V	I	I	I	●	I	●	●	I	I
T54A/S/V/G/C	T	T	●	●	●	●	●	●	●	●
V55A/F/I/K/T	V	V	●	●	●	●	●	●	●	●
D79E	D	D	●	●	●	**E**	●	●	●	●
Q80K/L/N/R/H/G	Q	Q	●	●	●	**G***	●	●	**K***	●
A87T	A	A	●	●	●	S	●	●	S	●
Y105C	Y	Y	●	●	●	●	●	**F**	●	●
R109K	R	R	●	●	●	●	●	●	●	●
R117H	R	R	●	●	●	●	●	●	●	●
S122G/A/R/N/T	S	S	●	●	●	**R**	●	T	T	T (<50%)
R123H/K	R	R	●	●	●	●	**T***	●	●	●
I132V	V	I	I	I	●	L	L	I	I	I
S138T/D/P	S	S	●	●	●	●	●	●	●	●
R155K/T/I/M/G/L/S/Q/P/N/W	R	R	●	●	●	●	●	●	●	●
A156S/T/V/I/F/N/G/D	A	A	●	●	●	●	●	●	●	●
V158I/M	V	V	●	●	●	●	●	●	●	●
V163L	V	V	●	●	●	●	●	●	●	●
D168Q/A/Y/V/E/T/N/P/I/H/G/F/S/K	D	D	●	●	●	●	**Q*?**	●	●	●
V170A/T/G/L/M	V	I	I	I	●	I	I	●	I	●
E173G	E	E	●	●	●	●	●	●	●	●
S174F/P	S	S	●	N	●	●	**T**	●	N	N
M175L	M	L	L	L*?	●	L	L*	L*	L*	●
E176G	E	E	●	●	●	**D**	**S**	●	●	●

[a] Symbols: ●, amino acid identical to the Con1 prototype replicon (HCV-1b); *, amino acid present in 100% of the sequences for the given HCV genotype; *?, amino acid present in 100% of the sequences for the given HCV genotype, with the exception of ambiguities. Amino acids without asterisks are present in the majority, but not all, of the sequences for the given HCV genotype. Amino acids unique for the given HCV genotype are presented in bold.

Y93H/C, GTs 1a and 1b showed a low genetic barrier in line with published clinical data (41), together with GT2, GT3, and GT4 isolates (score = 1). In contrast, GT5 and GT6 isolates showed scores of ≥5. Finally, for NS5B-M414, the genetic barrier was high only for GT2 isolates (score = 5) and low for GTs 1a, 1b, 3, 5, and 6 (score = 1). There were no differences in genetic barriers among HCV genotypes/subtypes for NS5B-S282 (all scores = 2.5).

Evolutionary history of natural RAVs. To find out whether naturally occurring RAVs are associated with ongoing random variation and/or founder effects in transmission clusters, the evolutionary history of these RAVs was traced by means of parsimony analyses, as implemented in Mesquite, using the phylogenetic trees obtained for each HCV variant and genomic region analyzed. Most natural RAVs seem to have occurred independently, although some grouped in well-supported clades (Fig. 1). The most significant cases (Fig. 1) are NS3-80K (GTs 1a and 6), NS5A-54H (GT1b) and 153L (GT1b), and NS5B-316N (GT1b) and 392I (GT2). Other RAVs forming well-supported clades are shown in Fig. S1 in the supplemental material: NS3-122G/T (GT1b) and 48V (GT6); NS5A-28M (GT1b), 31M (GT3), and 153L (GT3); and NS5B-414G/T (GT4). The same tree topologies were obtained when phylogenies were reconstructed from alignments without the codon positions implicated in resistance to DAAs (data not shown).

DISCUSSION

This study provides a comprehensive view on naturally occurring RAVs in all sequences publicly available to date for HCV genotypes 1 to 6. We obtained 2,901, 2,216, and 1,344 HCV sequences from HCV isolates from DAA-naive patients for NS3, NS5A, and NS5B, respectively, and they were screened for 28 NS3, 17 NS5A, and 58 NS5B relevant RAVs to compare their relative frequencies in the different HCV genotypes/subtypes. In addition, for the 1,126 HCV isolates in which sequences from the three genomic regions were available, we estimated the frequency of haplotypes carrying multiple RAVs to the four different classes of approved DAAs. Hence, to our knowledge, the current study represents the most comprehensive data set and analysis of naturally occurring HCV resistance to DAAs.

One of the most important findings is the remarkable difference observed in the prevalence of natural RAVs among HCV genotypes, with some RAVs being found in all isolates of a given genotype. First, for NS3, we found differences between genotypes for positions 36, 80, 117, 122, 168, and 175, in line with our previous results with a smaller data set (13) and with those observed by others (15, 16, 21, 22). Second, for NS5A, we identified relevant differences in the distribution of RAVs at positions 28, 30, 31, 58, 62, 93, and 153, particularly between GT1a and GT1b. Variation

TABLE 6 Majority amino acids and signatures of HCV genotypes 1 to 6 at sites associated with resistance to NS5a inhibitors

NS5a position	Amino acid for sequence or genotype[a]: Con1 reference (1b)	GT1 consensus amino acid	All GT1	1a	1b	2	3	4	5	6
L23F	L	L	●	●	●	●	●	●	●	●
Q24L	Q	K	K	K	●	S	S	K*	●	K
L28M/T	L	M	M	M	●	●	M	●	●	**V (<50%)**
R30Q/L	R	Q	Q	Q	●	**K**	**A**	●	Q	**S (<50%)**
L31M/V/F	L	L	●	●	●	M	●	M	●	●
P32L	P	P	●	●	●	●	●	●	●	●
Q54H/N/L/Y	Q	H	H	H	●	**T**	S	H*	S	H
P58S/A/L/T/H/D/P	P	H	H	H (49.5%)	**H**	●	●	●	●	●
Q62E/R/A/P/S	Q	E	E	E	●	**N**	**S**	E	**T**	**V**
A92T	A	A	●	●	●	**C**	**E**	●	●	●
Y93H/N/C	Y	Y	●	●	●	●	●	●	T*	T
F149L	F	F	●	●	●	●	●	●	●	●
V153M/L/I	V	L	L	L	●	●	●	●	L	●
M202L	M	M	●	●	●	●	●	●	●	●
M265V/T	M	M	●	●	●	●	●	●	●	●
D320E	D	D	●	●	●	●	●	●	**G***	●
Y321N	Y	Y	●	●	●	●	●	●	●	●

[a] Symbols: ●, amino acid identical to the Con1 prototype replicon (HCV-1b); *, amino acid present in 100% of the sequences for the given HCV genotype. Amino acids without asterisks are present in the majority, but not all, of the sequences for the given HCV genotype. Amino acids unique for the given HCV genotype are presented in bold.

in NS5A has been previously analyzed in 31 GT1a and 30 GT1b clinical isolates (17), finding similar frequencies for L28V in GT1a and L28M/V in GT1b. However, we report here a much lower prevalence of L31M in worldwide GT1a isolates (0.93%) and a higher or lower prevalence of Y93H/C in GT1a or GT1b isolates (0.70% versus 5.06%, respectively). Furthermore, while Paolucci et al. (17) did not find natural RAVs for NS5A inhibitors, we found up to 5.1% of HCV-1b isolates harboring RAVs. Third, NS5B RAVs to NNIs were frequent in positions 316, 392, 414, and 556 for certain HCV genotypes. For positions NS5B-316 and 556, our results are similar to those obtained previously in a smaller data set (20). However, we identified higher frequencies of RAVs at NS5B-414 in GT2 and GT4 isolates, probably because we considered M414Q/L (highly prevalent in these genotypes) a RAV. Other previous studies, analyzing a very limited number of isolates, also reported M414Q as a very frequent polymorphism in HCV GTs 2 and 3 (42). Finally, the most clinically relevant RAVs to NIs (L159F and S282T) were very infrequent (4.80% of GT1b isolates and 0.17%, 0.24%, and 1.24% of isolates from GTs 1a, 3, and 4, respectively). This near-absence of intergenotypic variation of NS5B-S282 suggests that NIs may indeed warrant their efficacy across genotypes, but the natural presence of the compensatory RAV L159F in a small proportion of GT1b isolates deserves further investigation. For other HCV genotypes,

TABLE 7 Majority amino acids and signatures of HCV genotypes 1 to 6 at sites associated with resistance to NS5B NIs

NS5B NI resistance position	Amino acid for sequence or genotype[a]: Con1 reference (1b)	GT1 consensus amino acid	All GT1	1a	1b	2	3	4	5	6
A15G	A	A	●	●	●	**G**	S	●	S*	●
K72M	K	K	●	●	●	●	●	●	●	●
S96T	S	S	●	●	●	●	●	●	●	●
N142T	N	N	●	●	●	●	●	●	●	●
L159F	L	L	●	●	●	●	●	●	●	●
R222Q	R	R	●	●	●	●	●	●	●	●
C223H/Y	C	C	●	●	●	●	●	●	●	●
I239V/L	I	I	●	●	●	●	●	**V**	●	●
S282T/R	S	S	●	●	●	●	●	●	●	●
A300T	A	S (<30%)	S (<30%)	Q	**S**	L	T	T	L*	Q
L320F/I	L	L	●	●	●	●	●	●	●	●
V321I	V	V	●	●	●	●	●	●	●	●
A396G	A	A	●	●	●	●	●	●	●	●
Y586C	Y	Y	●	●	●	F	F	F*#	F*#	F

[a] Symbols: ●, amino acid identical to the Con1 prototype replicon (HCV-1b); *, amino acid present in 100% of the sequences for the given HCV genotype; *#, amino acid present in 100% of the sequences for the given HCV genotype, with the exception of gaps. Amino acids without asterisks are present in the majority, but not all, of the sequences for the given HCV genotype. Amino acids unique for the given HCV genotype are presented in bold.

TABLE 8 Majority amino acids and signatures of HCV genotypes 1 to 6 at sites associated with resistance to NS5B NNIs

NS5B NNI resistance position	Amino acid for sequence or genotype[a]: Con1 reference (1b)	GT1 consensus amino acid	All GT1	1a	1b	2	3	4	5	6
T19S/P	T	Q	Q	**Q**	S	E	E*	S	E*	E*
K50R	K	K	●	●	●	●	●	●	●	●
D55E	D	D	●	●	●	●	●	●	●	●
M71V	M	V	V	V	●	V	V	I	●	I
H95Q/R	H	H	●	●	●	●	●	●	●	●
138I	I	I	●	●	●	●	●	●	●	●
L314F	L	L	●	●	●	●	●	●	●	●
C316Y/F/N/S	C	C	●	●	●	●	●	●	●	●
A338V	A	A	●	●	**V**	●	●	●	●	●
I363V	I	I	●	●	●	●	●	●	**V***	●
S365T/A/L/O/F	S	S	●	●	●	●	●	●	●	●
S368A/T	S	S	●	●	●	●	●	●	●	●
T389S/A	T	T	●	●	●	●	●	**E***	**Q***	●
L392I	L	L	●	●	●	**I**	●	●	●	●
N411S	N	N	●	●	●	●	●	●	●	●
M414L/T/I/V/Q	M	M	●	●	●	**Q**	●	**L**	●	●
L419M/V/S/I	L	L	●	●	●	I	I	I*?	●	I
A421V	A	A	●	●	●	V	V	V	●	V
R422K	R	R	●	●	●	●	●	●	●	●
M423T/V/I/A	M	M	●	●	●	●	●	●	**I***	●
I424V	I	I	●	●	●	V*	V*	V	V*	V
M426T/V/I	M	M	●	●	●	●	●	●	●	●
A442T	A	A	●	●	●	**N**	**P**	●	●	●
C445F	C	C	●	●	●	F*	F*	F*	F*	F*
I447F	I	I	●	●	●	M	M*	M*	M*	●
Y448H/C	Y	Y	●	●	●	●	●	●	●	●
C451R	C	C	●	●	●	V	T	T*	V*	T*
Y452H	Y	Y	●	●	●	●	●	●	●	●
C455F	E	E	●	●	●	**N**	T*	T*	T*	T
M/I462T	I	I	●	●	●	●	●	●	●	●
R465G	R	R	●	●	●	●	●	●	●	●
I482L/V/T/S	I	I	●	●	●	L*#	L*#	L*#	●	L*#
A486V/I/T/M	A	A	●	●	●	●	●	●	●	**G**
V494A/I	V	V	●	●	●	A*#	**C**	●	●	A
P495L/A/S/T/Q	P	P	●	●	●	●	●	●	●	●
P496A/T/S	P	P	●	●	●	●	●	●	●	●
V499A	V	A	A	A	●	A	A	A	A	A*#
R531K	R	R	●	●	●	K	●	K*#	K*	K
G554D/S	G	G	●	●	●	●	●	●	●	●
Y555C	Y	Y	●	●	●	A	**V**	A	A*	●
S556G/N/C	S	S	●	●	●	G	G	G	G*	●
G558R	G	G	●	●	●	●	**N**	●	●	●
D559G/S/N	D	D	●	●	●	●	●	●	●	●
W571R	W	W	●	●	●	L	**H**	Y	Y	**M (<50%)**

[a] Symbols: ●, amino acid identical to the Con1 prototype replicon (HCV-1b); *, amino acid present in 100% of the sequences for the given HCV genotype; *#, amino acid present in 100% of the sequences for the given HCV genotype, with the exception of gaps; *?, amino acid present in 100% of the sequences for the given HCV genotype, with the exception of ambiguities. Amino acids without asterisks are present in the majority, but not all, of the sequences for the given HCV genotype. Amino acids unique for the given HCV genotype are presented in bold.

NS5B RAVs identified in HCV replicon assays (A15G and I239V) showed a significant prevalence in GT2 and GTs 4 and 6, respectively, but their potential clinical significance is unclear.

A second important finding is the presence of multiple RAVs to several DAAs in HCV isolates from DAA-naive patients, irrespective of the viral genotype. Considering those haplotypes with more than one RAV, the most frequent for GT1a showed at least two RAVs at NS3, but for the rest of the HCV genotypes, finding three or more RAVs involving at least two different genes was very common. These results indicate intergenotypic differences in the frequency and combinations of natural RAVs that may have potential clinical implications. However, the real levels of resistance conferred by these natural RAVs should be determined with specific assays for each genotype/subtype and not only with HCV-1 replicons (see below). Despite such differences, the most common

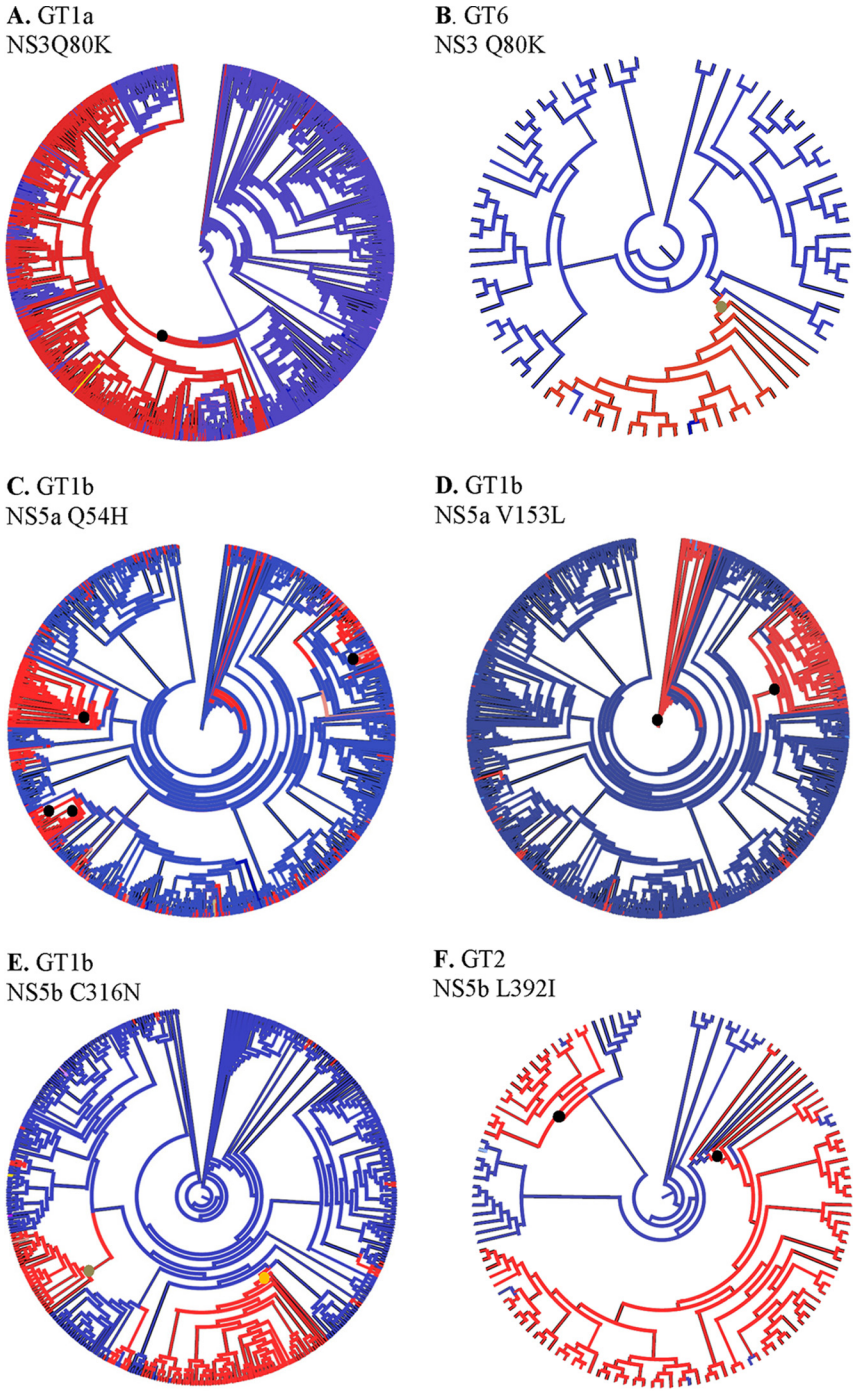

FIG 1 Dendrograms representing the evolutionary history of the six most representative resistance mutations forming well-supported clades (A to F). Blue lines represent nonresistant lineages; red lines represent lineages carrying the resistance mutation; black dots represent the emergence of well-supported clades of resistance (aLRT, >0.85). In panel B, the node including the HCV-6a clade is indicated in green. In panel E, nodes including the Asian (yellow) and U.S. (green) clades are indicated.

combinations of natural RAVs included NS3-36L+175L and/or 80G/K (GTs 1a, 2, 3, 4, and 5), NS3-80G/K+122G/R/T/N (GTs 2, 5, and 6), and NS3-122T/R+NS5B-556G and/or 175L (GTs 2, 4, and 5). While the selection of several combinations of RAVs during DAA treatment has been reported previously for TPV (NS3-36M+155K), SMV (NS3-36M+155K; 80L+155K; 122R+155K; 155K+168E; 80R+168E), and MCB (NS5B-159F+320F) (38), none of these combinations was found in any of the HCV isolates from DAA-naive patients analyzed here. Collectively, these observations suggest that natural RAV profiles may be different from those induced during treatment with DAAs.

The third important finding of our study is that several

RAVs previously characterized for HCV GT1 (either in replicon assays or during DAA treatment failure) were present as majority or signature (unique) amino acids for other viral genotypes (detailed in Tables 5 to 8; see also Table S4 in the supplemental material). Further, we found several positions with differential signature/majority amino acid patterns between subtypes of the same GTs. The potential impact of these genotype/subtype-specific natural RAVs on the pan-genotypic efficacy of new interferon-free DAA regimens remains to be determined and deserves further research. An important limitation of our study is that the RAVs investigated here have been mainly characterized for HCV GT1 and might not necessarily confer resistance in other HCV genotypes. Previous work reported reduced DAA sensitivity in isolates from non-1 HCV genotypes carrying some of the consensus and/or signature amino acids identified here. Clear examples are NS3-Q80K for GT1a, NS3-Q80G for GT2, and NS3-D168Q for GT3 (13, 16, 43). However, this may not always be the case: Y93H confers several-thousandfold resistance to NS5A inhibitors in GT1 isolates, but GT5 and GT6 isolates with the naturally occurring variant Y93T do not show a great degree of natural resistance, and isolates with T93H show only low levels of resistance (44). Because the information available on specific RAVs for genotype 2 to 6 isolates (either *in vitro* in replicon assays or *in vivo*) is still limited, this issue remains beyond the scope of our study.

Nevertheless, we were able to identify several RAVs in non-GT1 isolates that are known to confer resistance in chimeric HCV replicons for the corresponding genotype (see Table S1 in the supplemental material), such as NS3-D/Q168E for GTs 5 and 6, NS5A-L28M/F/V for GTs 2 and 4, L/A30K/H and Y93H for GT3, and NS5B-L329I and V494 for GT2. Our results strongly suggest that DAA resistance profiles may be helpful for the choice of the most appropriate DAA combination for each particular HCV genotype/subtype. Thus, the identification of RAV profiles for GTs 2 to 6 *in vitro* using genotype/subtype-specific assays is urgently needed, together with their correlation with *in vivo* data from the increasing number of patients infected with HCV GTs 2 to 6 being treated with DAAs.

Given the high rates of evolution of HCV, the development of resistance during treatment is inevitable. Strikingly, our results show that most HCV isolates have a low genetic barrier toward mutations to generate RAVs to different DAAs, regardless of the HCV genotype. These results may seem difficult to reconcile with the high efficacy of DAA combinations but are in line with the clinical observations that SVR rates have been increasing when more DAAs are being included in treatment regimens. All the studies on DAAs administered as monotherapy showed the rapid selection of RAVs (45). While the combination of one DAA with P/R marked a milestone in achieving SVR rates higher than 70%, not until the use of two or more DAAs in combination did SVR rates reach 85 to 90%, presumably because of the continuous selection of RAVs in amino acid sites with low genetic barriers in single-DAA regimens. While this may hold true for PIs and NNIs, it is probably not that relevant for combinations with highly potent NS5A inhibitors and/or those with high clinical genetic barriers (NIs such as SOF or MCB). During treatment with these highly potent DAAs that immediately reduce viral replication, the virus probably has no capacity to generate and allow the spread of new RAVs emerging as minor variants in the viral population. However, a different question is whether DAA escape occurs because a natural RAV preexists as a predominant variant within an individual. Despite the high natural variability of HCV, SVR rates observed in pivotal studies are much higher than 90% for HCV GT1, but as different IFN-free DAA combinations are generalized for treatment of non-1 HCV genotypes, natural resistance/susceptibility profiling may still be useful for choosing the best drug regimens and for retreatment of relapses.

Finally, from a public health perspective it is important to investigate whether naturally occurring RAVs appear randomly without antiviral pressure, correspond to old transmission clusters, and/or share a geographical distribution pattern. Using phylogenetic analyses, we revealed that most natural RAVs appeared *de novo*, as they are found in external branches. However, in some cases such mutations were clearly differentiated in well-supported clades. It has been suggested recently that the majority of GT1a infections carrying the 80K variant (associated with resistance to SMV) were transmitted from a single origin (21). Our analyses further support their results and shed more light on the evolution of NS3-Q80K. We found that this variant in GT6 occurs only in subtype 6a, being present in 18 out of 20 of all GT-6a sequences, 13 of them from Hong Kong. In addition, we identified other well-differentiated clades grouping different RAVs, NS5A-153L and NS5B-316N in GT1b isolates. It is worth noting that our phylogenetic analyses support the existence of two distinct GT1b clades carrying NS5B-316N: one clade from Asian isolates (Japanese and Chinese sequences) and a second one from the United States. For these highly prevalent natural RAVs, surveillance may be useful to asses their potential impact on the efficacy of some of the new IFN-free DAA regimens containing NNIs. In addition, if proven clinically relevant, these natural RAVs also deserve attention for their potential to be transmitted in small local epidemics associated with risk behaviors. In addition, and despite the high efficacy of new DAA regimens shown in pivotal studies, in actual clinical practice a small but significant number of patients may fail DAA treatment and potentially transmit selected DAA-resistant HCV variants, which need to be identified and characterized.

In conclusion, our comprehensive analysis of natural HCV polymorphisms associated with resistance to DAAs indicates that natural RAVs are relatively common in viral isolates from treatment-naive patients, with frequencies clearly varying among HCV genotypes/subtypes. Importantly, for any HCV genotype, there is a significant proportion of viral isolates carrying at least two RAVs, and we identified distinct genotypic profiles for HCV natural resistance to DAAs. Furthermore, we observed that, for any HCV genotype, most of the more prevalent wild-type codons present low genetic barriers for the generation of new RAVs, emphasizing the importance of using at least one highly potent drug as a backbone in IFN-free DAA combinations. While our evolutionary analyses showed that the majority of naturally occurring RAVs are probably being generated at random and very recently, a few natural RAVs are associated with long-lasting, geographically delimited successful transmission with a high potential to keep being further transmitted over time.

FUNDING INFORMATION

Ministerio de Economía y Competitividad (MINECO) | Instituto de Salud Carlos III (ISCIII) provided funding to F. Xavier Lopez-Labrador under grant number PI10/00512. Ministerio de Economía y Competitividad (MINECO) | Instituto de Salud Carlos III (ISCIII) provided funding to F. Xavier Lopez-Labrador under grant number PI10/01734. Ministerio de Economía y Competitividad (MINECO) provided funding to Fernando González-Candelas under grant number BFU2011-24112. Ministerio de Economía y Competitividad (MINECO) provided funding to Fernando González-Candelas under grant number BFU2014-58656R. Generalitat Valenciana (Regional Government of Valencia) provided funding to F. Xavier Lopez-Labrador under grant number FPA/2013/A/083. Generalitat Valenciana (Regional Government of Valencia) provided funding to F. Xavier Lopez-Labrador under grant number ACOMP/2013/086. Ministerio de Educación, Cultura y Deporte (MECD) provided funding to Juan Angel Patiño-Galindo under grant number FPU-AP2010-0561.

REFERENCES

1. **Hajarizadeh B, Grebely J, Dore GJ.** 2013. Epidemiology and natural history of HCV infection. Nat Rev Gastroenterol Hepatol **10:**553–562. http://dx.doi.org/10.1038/nrgastro.2013.107.
2. **Mohd Hanafiah K, Groeger J, Flaxman AD, Wiersma ST.** 2013. Global epidemiology of hepatitis C virus infection: new estimates of age-specific antibody to HCV seroprevalence. Hepatology **57:**1333–1342. http://dx.doi.org/10.1002/hep.26141.
3. **Smith DB, Bukh J, Kuiken C, Muerhoff AS, Rice CM, Stapleton JT, Simmonds P.** 2014. Expanded classification of hepatitis C virus into 7 genotypes and 67 subtypes: updated criteria and genotype assignment web resource. Hepatology **59:**318–327. http://dx.doi.org/10.1002/hep.26744.
4. **Bukh J.** 1995. Genetic heterogeneity of hepatitis C virus: quasispecies and genotypes. Semin Liver Dis **15:**41–63. http://dx.doi.org/10.1055/s-2007-1007262.
5. **Okamoto H, Kojima M, Okada S, Yoshizwa H, Iizuka H, Tanaka T, Muchmore EE, Peterson DA, Ito Y, Mishiro S.** 1992. Genetic drift of hepatitis C virus during an 8.2-year infection in a chimpanzee: variability and stability. Virology **190:**894–899. http://dx.doi.org/10.1016/0042-6822(92)90933-G.
6. **Mangia A, Santoro R, Minerva N, Ricci GL, Caretta V, Persico M, Vinelli F, Scotto G, Bacca D, Annese M, Romano M, Zechini F, Sogari F, Spirito F, Andriulli A.** 2005. Peginterferon alfa-2b and ribavirin for 12 vs. 24 weeks in HCV genotype 2 or 3. N Engl J Med **352:**2609–2617. http://dx.doi.org/10.1056/NEJMoa042608.
7. **Pang PS, Planet PJ, Glenn JS.** 2009. The evolution of the major hepatitis C genotypes correlates with clinical response to interferon therapy. PLoS One **4:**e6579. http://dx.doi.org/10.1371/journal.pone.0006579.
8. **Zein NN.** 2000. Clinical significance of hepatitis C virus genotypes. Clin Microbiol Rev **13:**223–235. http://dx.doi.org/10.1128/CMR.13.2.223-235.2000.
9. **AASLD/ISDA HCV Guidance Panel.** 2015. Hepatitis C guidance: AASLD-ISDA recommendations for testing, managing, and treating adults infected with hepatitis C virus. Hepatology **62:**932–954. http://dx.doi.org/10.1002/hep.27950.
10. **European Association for the Study of the Liver.** 2015. EASL recommendations on treatment of hepatitis C 2015. J Hepatol **63:**199–236. http://dx.doi.org/10.1016/j.jhep.2015.03.025.
11. **Bartenschlager R, Lohmann V, Penin F.** 2013. The molecular and structural basis of advanced antiviral therapy for hepatitis C virus infection. Nat Rev Microbiol **11:**482–496. http://dx.doi.org/10.1038/nrmicro3046.
12. **Asselah T, Boyer N, Saadoun D, Martinot-Peignoux M, Marcellin P.** 2016. Direct-acting antivirals for the treatment of hepatitis C virus infection: optimizing current IFN-free treatment and future perspectives. Liver Int **36:**47–57. http://dx.doi.org/10.1111/liv.13027.
13. **López-Labrador FX, Moya A, González-Candelas F.** 2008. Mapping natural polymorphisms of hepatitis C virus NS3/4A protease and antiviral resistance to inhibitors in worldwide isolates. Antivir Ther **13:**481–494.
14. **Bartels DJ, Zhou Y, Zhang EZ, Marcial M, Byrn RA, Pfeiffer T, Tigges AM, Adiwijaya BS, Lin C, Kwong AD, Kieffer TL.** 2008. Natural prevalence of hepatitis C virus variants with decreased sensitivity to NS3-4A protease inhibitors in treatment-naive subjects. J Infect Dis **198:**800–807. http://dx.doi.org/10.1086/591141.
15. **Paolucci S, Fiorina L, Piralla A, Gulminetti R, Novati S, Barbarini G, Sacchi P, Gatti M, Dossena L, Baldanti F.** 2012. Naturally occurring mutations to HCV protease inhibitors in treatment-naïve patients. Virol J **9:**245. http://dx.doi.org/10.1186/1743-422X-9-245.
16. **Alves R, Queiro AT, Pessoa MG, da Silva EF, Mazo DF, Carrilho FJ, Carvalho-Filho RJ, de Carvalho IM.** 2013. The presence of resistance mutations to protease and polymerase inhibitors in hepatitis C virus sequences from the Los Alamos databank. J Viral Hepat **20:**414–421. http://dx.doi.org/10.1111/jvh.12051.
17. **Paolucci S, Fiorina L, Mariani B, Gulminetti R, Novati S, Barbarini G, Bruno R, Baldanti F.** 2013. Naturally occurring resistance mutations to inhibitors of HCV NS5A region and NS5B polymerase in DAA treatment-naïve patients. Virol J **10:**355. http://dx.doi.org/10.1186/1743-422X-10-355.
18. **de Carvalho IM, Alves R, de Souza PA, da Silva EF, Mazo D, Carrilho FJ, Queiroz AT, Pessoa MG.** 2014. Protease inhibitor resistance mutations in untreated Brazilian patients infected with HCV: novel insights about targeted genotyping approaches. J Med Virol **86:**1714–1721. http://dx.doi.org/10.1002/jmv.24015.
19. **Margeridon-Thermet S, Le Pogam S, Li L, Liu TF, Shulman N, Shafer RW, Najera I.** 2014. Similar prevalence of low-abundance drug-resistant variants in treatment-naive patients with genotype 1a and 1b hepatitis C virus infections as determined by ultradeep pyrosequencing. PLoS One **9:**e105569. http://dx.doi.org/10.1371/journal.pone.0105569.
20. **Di Maio VC, Cento V, Mirabelli C, Artese A, Costa G, Alcaro S, Perno CF, Ceccherini-Silberstein F.** 2014. Hepatitis C virus genetic variability and the presence of NS5B resistance-associated mutations as natural polymorphisms in selected genotypes could affect the response to NS5B inhibitors. Antimicrob Agents Chemother **58:**2781–2797. http://dx.doi.org/10.1128/AAC.02386-13.
21. **McCloskey RM, Liang RH, Joy JB, Krajden M, Montaner JS, Harrigan PR, Poon AF.** 2015. Global origin and transmission of hepatitis C virus nonstructural protein 3 Q80K polymorphism. J Infect Dis **211:**1288–1295. http://dx.doi.org/10.1093/infdis/jiu613.
22. **Cento V, Mirabelli C, Salpini R, Dimonte S, Artese A, Costa G, Mercurio F, Svicher V, Parrotta L, Bertoli A, Ciotti M, Di Paolo D, Sarrecchia C, Andreoni M, Alcaro S, Angelico M, Perno CF, Ceccherini-Silberstein F.** 2012. HCV genotypes are differently prone to the development of resistance to linear and macrocyclic protease inhibitors. PLoS One **7:**e39652. http://dx.doi.org/10.1371/journal.pone.0039652.
23. **Combet C, Garnier N, Charavay C, Grando D, Crisan D, Lopez J, Dehne-Garcia A, Geourjon C, Bettler E, Hulo C, Le Mercier P, Bartenschlager R, Diepolder H, Moradpour D, Pawlotsky JM, Rice CM, Trepo C, Penin F, Deleage G.** 2007. The European hepatitis C virus database. Nucleic Acids Res **35**(Database Issue)**:**D363–D366.
24. **Kuiken C, Yusim K, Boykin L, Richardson R.** 2005. The Los Alamos HCV sequence database. Bioinformatics **21:**379–384. http://dx.doi.org/10.1093/bioinformatics/bth485.
25. **Pickett B, Greer D, Zhang Y, Stewart L, Zhou L, Sun G, Gu Z, Kumar S, Zaremba S, Larsen C, Jen W, Klem E, Scheuermann R.** 2012. Virus Pathogen Database and Analysis Resource (ViPR): a comprehensive bioinformatics database and analysis resource for the coronavirus research community. Viruses **4:**3209–3226. http://dx.doi.org/10.3390/v4113209.
26. **Edgar RC.** 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics **5:**113. http://dx.doi.org/10.1186/1471-2105-5-113.
27. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32:**1792–1797. http://dx.doi.org/10.1093/nar/gkh340.
28. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S.** 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony. Methods Mol Biol Evol **28:**2731–2739. http://dx.doi.org/10.1093/molbev/msr121.
29. **Huang Y, Niu B, Gao Y, Fu L, Li W.** 2010. CD-HIT Suite: a Web server for clustering and comparing biological sequences. Bioinformatics **26:**680–682. http://dx.doi.org/10.1093/bioinformatics/btq003.
30. **Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O.** 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol **59:**307–321. http://dx.doi.org/10.1093/sysbio/syq010.
31. **Salvatierra KA.** 2014. Ph.D. thesis. University of Valencia, Valencia, Spain.
32. **Hall TA.** 1999. BioEdit: a user-friendly biological sequence alignment

editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser **41:**95–98.
33. **R Core Team.** 2014. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.
34. **Korber B, Myers G.** 1992. Signature pattern analysis: a method for assessing viral sequence relatedness. AIDS Res Hum Retroviruses **8:**1549–1560. http://dx.doi.org/10.1089/aid.1992.8.1549.
35. **Excoffier L, Lischer HE.** 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour **10:**564–567. http://dx.doi.org/10.1111/j.1755-0998.2010.02847.x.
36. **Van de Vijver DA, Wensing AM, Angarano G, Asjö B, Balotta C, Boeri E, Camacho R, Chaix ML, Costagliola D, De Luca A, Derdelinckx I, Grossman Z, Hamouda O, Hatzakis A, Hemmer R, Hoepelman A, Horban A, Korn K, Kücherer C, Leitner T, Loveday C, MacRae E, Maljkovic I, de Mendoza C, Meyer L, Nielsen C, Op de Coul EL, Ormaasen V, Paraskevis D, Perrin L, Puchhammer-Stöckl E, Ruiz L, Salminen M, Schmit JC, Schneider F, Schuurman R, Soriano V, Stanczak G, Stanojevic M, Vandamme AM, Van Laethem K, Violin M, Wilbe K, Yerly S, Zazzi M, Boucher CA.** 2006. The calculated genetic barrier for antiretroviral drug resistance substitutions is largely similar for different HIV-1 subtypes. J Acquir Immune Defic Syndr **41:**352–360. http://dx.doi.org/10.1097/01.qai.0000209899.05126.e4.
37. **Svicher V, Cento V, Salpini R, Mercurio F, Fraune M, Beggel B, Han Y, Gori C, Wittkop L, Bertoli A, Micheli V, Gubertini G, Longo R, Romano S, Visca M, Gallinaro V, Marino N, Mazzotta F, De Sanctis GM, Fleury H, Trimoulet P, Angelico M, Cappiello G, Zhang XX, Verheyen J, Ceccherini-Silberstein F, Perno CF.** 2011. Role of hepatitis B virus genetic barrier in drug-resistance and immune-escape development. Dig Liver Dis **43:**975–983. http://dx.doi.org/10.1016/j.dld.2011.07.002.
38. **Wyles DL, Gutierrez JA.** 2014. Importance of HCV genotype 1 subtypes for drug resistance and response to therapy. J Viral Hepat **21:**229–240. http://dx.doi.org/10.1111/jvh.12230.
39. **Svarovskaia ES, Dvory-Sobol H, Gontcharova V, Chiu S, Hebner CM, Hyland R, Kowdley K, Lawitz E, Gane E, Symonds WT, McHutchison JG, Miller MD, Mo H.** 2012. Comprehensive resistance testing in patients who relapsed after treatment with sofosbuvir (GS-7977)-containing regimens in phase 2 studies. Hepatology **56**(Suppl 1)**:**551A.
40. **Tong X, Li L, Haines K, Najera I.** 2014. Identification of the NS5B S282T resistant variant and two novel amino acid substitutions that affect replication capacity in hepatitis C virus-infected patients treated with mericitabine and danoprevir. Antimicrob Agents Chemother **58:**3105–3114. http://dx.doi.org/10.1128/AAC.02672-13.
41. **McPhee F, Hernandez D, Yu F, Ueland J, Monikowski A, Carifa A, Falk P, Wang C, Fridell R, Eley T, Zhou N, Gardiner D.** 2013. Resistance analysis of hepatitis C virus genotype 1 prior treatment null responders receiving daclatasvir and asunaprevir. Hepatology **58:**902–911. http://dx.doi.org/10.1002/hep.26388.
42. **Legrand-Abravanel F, Henquell C, Le Guillou-Guillemette H, Balan V, Mirand A, Dubois M, Lunel-Fabiani F, Payan C, Izopet J.** 2009. Naturally occurring substitutions conferring resistance to hepatitis C virus polymerase inhibitors in treatment-naive patients infected with genotypes 1-5. Antivir Ther **14:**723–730.
43. **Thibeault D, Bousquet C, Gingras R, Lagacé L, Maurice R, White PW, Lamarre D.** 2004. Sensitivity of NS3 serine proteases from hepatitis C virus genotypes 2 and 3 to the inhibitor BILN 2061. J Virol **78:**7352–7359. http://dx.doi.org/10.1128/JVI.78.14.7352-7359.2004.
44. **Scheel TK, Gottwein JM, Mikkelsen LS, Jensen TB, Bukh J.** 2011. Recombinant HCV variants with NS5A from genotypes 1-7 have different sensitivities to an NS5A inhibitor but not interferon-α. Gastroenterology **140:**1032–1042. http://dx.doi.org/10.1053/j.gastro.2010.11.036.
45. **Sarrazin C, Zeuzem S.** 2010. Resistance to direct antiviral agents in patients with hepatitis C virus infection. Gastroenterology **138:**447–462. http://dx.doi.org/10.1053/j.gastro.2009.11.055.